METHODS AND MEANS FOR PRODUCING EFFICIENT SILENCING CONSTRUCT USING RECOMBINATIONAL CLONING

[001] CROSS-REFERENCE TO RELATED APPLICATIONS

[002] This application is a continuation-in-part of U.S. Provisional Application Serial No. 60/264,067, filed January 25, 2001, and a continuation-in-part of U.S. Provisional Application Serial No. 60/333,743, filed November 29, 2001. The contents of these applications are incorporated herein by reference.

[003] FIELD OF THE INVENTION

[004] This invention relates to efficient methods and means for producing chimeric nucleic acid constructs capable of producing dsRNA useful for silencing target nucleic acid sequences of interest. The efficiency of the disclosed methods and means further allows high throughput analysis methods to determine the function of isolated nucleic acids, such as ESTs, without a known function and may further be put to use to isolate particular genes or nucleotide sequences from a preselected group of genes.

[005] BACKGROUND ART

[006] Increasingly, the nucleotide sequence of whole genomes of organisms, including *Arabidopsis thaliana*, has been determined and as these data become available, they provide a wealth of unmined information. The ultimate goal of these genome projects is to identify the biological function of every gene in the genome.

[007] Attribution of a function to a nucleic acid with a particular nucleotide sequence can be achieved in a variety of ways. Some of the genes have been characterized directly using the appropriate assays. Others have been attributed with a tentative function through homology with (parts of) genes having a known function in other organisms. Loss-of-function mutants, obtained e.g. by tagged insertional mutagenesis have also been very informative about the role of some

of these unknown genes (AzpiroLeehan and Feldmann 1997; Martienssen 1998) particularly in the large-scale analysis of the yeast genome (Ross-MacDonald et al., 1999).

[008] Structural mutants resulting in a loss-of-function may also be mimicked by interfering with the expression of a nucleic acid of interest at the transcriptional or post-transcriptional level. Silencing of genes, particularly plant genes using anti-sense or co-suppression constructs to identify gene function, especially for a larger number of targets, is however hampered by the relatively low proportion of silenced individuals obtained, particularly those wherein the silencing level is almost complete.

[009] Recent work has demonstrated that the silencing efficiency could be greatly improved both on quantitative and qualitative level using chimeric constructs encoding RNA capable of forming a double stranded RNA by basepairing between the antisense and sense RNA nucleotide sequences respectively complementary and homologous to the target sequences.

[010] Fire *et al.*, 1998 describe specific genetic interference by experimental introduction of double-stranded RNA in *Caenorhabditis elegans.* The importance of these findings for functional genomics has been discussed (Wagner and Sun, 1998).

[011] WO 99/32619 provides a process of introducing RNA into a living cell to inhibit gene expression of a target gene in that cell. The process may be practiced ex vivo or in vivo. The RNA has a region with double-stranded structure. Inhibition is sequence-specific in that the nucleotide sequences of the duplex region of the RNA and or a portion of the target gene are identical.

[012] Waterhouse *et al.* 1998 describes that virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and anti-sense RNA. The sense and antisense RNA may be located in one transcript that has self-complementarity.

[013] Hamilton *et al.* 1998 describes that a transgene with repeated DNA, i.e. inverted copies of its 5' untranslated region, causes high frequency, post-transcriptional suppression of ACC-oxidase expression in tomato.

[014] WO 98/53083 describes constructs and methods for enhancing the inhibition of a target gene within an organism, which involve inserting into the gene-silencing vector an inverted, repeat sequence of all or part of a polynucleotide region within the vector.

[015] WO 99/53050 provides methods and means for reducing the phenotypic expression of a nucleic acid of interest in eukaryotic cells, particularly in plant cells. These methods involve introducing chimeric genes encoding sense and antisense RNA molecules directed towards the target nucleic acid, which are capable of forming a double stranded RNA region by base-pairing between the regions with the sense and antisense nucleotide sequence, or introducing the RNA molecules themselves. Preferably, the RNA molecules comprise simultaneously both sense and antisense nucleotide sequences.

[016] WO 99/49029 relates generally to a method of modifying gene expression and to synthetic genes for modifying endogenous gene expression in a cell, tissue or organ of a transgenic organism, in particular to a transgenic animal of plant. Synthetic genes and genetic constructs, capable of forming a dsRNA which are capable of repressing, delaying or otherwise reducing the expression of an endogenous gene or a target gene in an organism when introduced thereto are also provided.

[017] WO 99/61631 relates to methods to alter the expression of a target gene in a plant using sense and antisense RNA fragments of the gene. The sense and antisense RNA fragments are capable of pairing and forming a double-stranded RNA molecule, thereby altering the expression of the gene. The present invention also relates to plants, their progeny and seeds thereof obtained using these methods.

[018] WO 00/01846 provides a method of identifying DNA responsible for conferring a particular phenotype in a cell. That method comprises a) constructing a cDNA or genomic library of the DNA of the cell in a suitable vector in an orientation relative to (a) promoter(s) capable of initiating transcription of the cDNA or DNA to double stranded (ds) RNA upon binding of an appropriate transcription factor to the promoter(s); b) introducing the library into one or

more of cells comprising the transcription factor, and c) identifying and isolating a particular phenotype of a cell comprising the library and identifying the DNA or cDNA fragment from the library responsible for conferring the phenotype. Using this technique, it is also possible to assign function to a known DNA sequence by a) identifying homologues of the DNA sequence in a cell, b) isolating the relevant DNA homologue(s) or a fragment thereof from the cell, c) cloning the homologue or fragment thereof into an appropriate vector in an orientation relative to a suitable promoter capable of initiating transcription of dsRNA from said DNA homologue or fragment upon binding of an appropriate transcription factor to the promoter and d) introducing the vector into the cell from step a) comprising the transcription factor.

[019] WO 00/44914 also describes composition and methods for in vivo and in vitro attenuation of gene expression using double stranded RNA, particularly in zebrafish.

[020] WO 00/49035 discloses a method for silencing the expression of an endogenous gene in a cell. That method involves overexpressing in the cell a nucleic acid molecule of the endogenous gene and an antisense molecule including a nucleic acid molecule complementary to the nucleic acid molecule of the endogenous gene, wherein the overexpression of the nucleic acid molecule of the endogenous gene and the antisense molecule in the cell silences the expression of the endogenous gene.

[021] Smith et al., 2000 as well as WO 99/53050 described that intron containing dsRNA further increased the efficiency of silencing.

[022] However, the prior art has not solved the problems associated with the efficient conversion of any nucleotide sequence of interest into a chimeric construct capable of producing a dsRNA in eukaryotic cells, particularly in plant cells, and preferably in a way amenable to the processing of large number of nucleotide sequences.

[023] These and other problems have been solved as described hereinafter in the different embodiments and claims.

[024] SUMMARY OF THE INVENTION

[025] It is an object of the invention to provide vectors comprising the following operably linked DNA fragments a) an origin of replication allowing replication in microorganisms (1), preferably bacteria; particularly *Escherichia coli*; b) a selectable marker region (2) capable of being expressed in microorganisms, preferably bacteria; and c) a chimeric DNA construct comprising in sequence (i) a promoter or promoter region (3) capable of being recognized by RNA polymerases of a eukaryotic cell, preferably a plant-expressible promoter; (ii) a first recombination site (4), a second recombination site (5), a third recombination site (6) and a fourth recombination site (7); and (iii) a 3' transcription terminating and polyadenylation region (8) functional in the eukaryotic cell; wherein the first recombination site (4) and the fourth recombination site (7) are capable of reacting with a same recombination site, preferably are identical, and the second recombination site (5) and the third recombination site (6), are capable of reacting with a same recombination site, preferably are identical; and wherein the first recombination site (4) and the second recombination site (5) do not recombine with each other or with a same recombination site or the third recombination site (6) and the fourth recombination site (7) do not recombine with each other or with a same recombination site. Optionally the vector may further include additional elements such as: a second selectable marker gene (9) between the first (4) and second recombination site (5) and/or a third selectable marker gene (10) between the third (6) and fourth recombination site (7) and/or a region flanked by intron processing signals (11), preferably an intron, functional in the eukaryotic cell, located between the second recombination site (5) and the third recombination site (6) and/or a fourth selectable marker gene (19), located between the second (5) and third recombination site (6) and/or left and right border T-DNA sequences flanking the chimeric DNA construct and/or a selectable marker gene capable of being expressed in eukaryotic , preferably plant, cells, preferably located between the left and the right T-DNA border sequences and/or an origin of replication capable of functioning in *Agrobacterium* spp. Selectable marker genes

may be selected from the group consisting of an antibiotic resistance gene, a tRNA gene, an auxotrophic marker, a toxic gene, a phenotypic marker, an antisense oligonucleotide; a restriction endonuclease; a restriction endonuclease cleavage site, an enzyme cleavage site, a protein binding site, an a sequence complementary PCR primer. Preferably the first (4) and fourth recombination site (7) are *att*R1 comprising the nucleotide sequence of SEQ ID No 4 and the second (5) and third (6) recombination site are *att*R2 comprising the nucleotide sequence of SEQ ID No 5 or the first (4) and fourth recombination site (7) are *att*P1 comprising the nucleotide sequence of SEQ ID No 10 and the second (5) and third (6) recombination site are *att*P2 comprising the nucleotide sequence of SEQ ID No 11.

[026] It is another objective of the invention to provide a kit comprising an acceptor vector according to invention, preferably further comprising at least one recombination protein capable of recombining a DNA segment comprising at least one of the recombination sites.

[027] It is yet another objective of the invention to provide a method for making a chimeric DNA construct capable of expressing a dsRNA in a eukaryotic cell comprising the steps of combining *in vitro*:

an acceptor vector as herein before described;

an insert DNA, preferably a linear or circular insert DNA, comprising

a DNA segment of interest (12) flanked by

a fifth recombination site (13) which is capable of recombining with the first (4) or fourth recombination site (7) on the vector; and

a sixth recombination site (14) which is capable of recombining with the second (5) or third recombination site (6) on the vector;

at least one site specific recombination protein capable of recombining the first (4) or fourth (7) and the fifth recombination site (13) and the second (5) or third (6) and the sixth recombination site (14);

allowing recombination to occur in the presence of at least one recombination protein, preferably selected from Int and IHF and (ii) Int, Xis, and IHF, so as to produce a reaction mixture comprising product DNA molecules, the product DNA molecule comprising in sequence:

the promoter or promoter region (3) capable of being recognized by RNA polymerases of the eukaryotic cell;

a recombination site (15) which is the recombination product of the first (4) and the fifth recombination site (13);

the DNA fragment of interest (12);

a recombination site (16) which is the recombination product of the second (4) and the sixth recombination site (14);

a recombination site (17) which is the recombination product of the third (5) and the sixth recombination site (14);

the DNA fragment of interest in opposite orientation (12);

a recombination site (18) which is the recombination product of the fourth (7) and the fifth recombination site (13); and

the 3' transcription terminating and polyadenylation region (8) functional in the eukaryotic cell; and

selecting the product DNA molecules, preferably *in vivo*.

[028] The method allows that multiple insert DNAs comprising different DNA fragments of interest are processed simultaneously.

[029] The invention also provides a method for preparing a eukaryotic non-human organism, preferably a plant, wherein the expression of a target nucleic acid of interest is reduced or inhibited, the method comprising:

preparing a chimeric DNA construct capable of expressing a dsRNA in cells of the eukaryotic non-human organism according to methods of the invention;

introducing the chimeric DNA construct in cells of the eukaryotic non-human organism; and

isolating the transgenic eukaryotic organism.

[030] It is also an objective of the invention to provide a method for isolating a nucleic acid molecule involved in determining a particular trait, comprising the steps of:

preparing a library of chimeric DNA constructs capable of expressing a dsRNA in cells of the eukaryotic non-human organism according to any one of the methods of the invention;

introducing individual representatives of the library of chimeric DNA constructs in cells of the eukaryotic non-human organism;

isolating a eukaryotic organism exhibiting the particular trait; and

isolating the nucleic acid molecule.

[031] The invention also provides a eukaryotic non-human organism, preferably a plant comprising a chimeric DNA construct obtainable through the methods of the invention.

[032] BRIEF DESCRIPTION OF THE FIGURES

[033] Fig 1. Schematic representation of vectors and method used in a preferred embodiment of the invention.

[034] Fig 1A: A nucleic acid of interest (12) is amplified by PCR using primers comprising two different recombination sites (13, 14) which cannot react with each other or with the same other recombination site. This results in "insert DNA" wherein the nucleic acid of interest (12) is flanked by two different recombination sites (13, 14).

[035] Fig 1B. Using at least one recombination protein, the insert DNA is allowed to recombine with the acceptor vector between the recombination sites, whereby the first (4) and fourth recombination site (7) react with one of the recombination sites (13) flanking the PCR amplified DNA of interest (12) and the second (5) and third (6) recombination site on the acceptor vector recombine with the other recombination site (14) flanking the DNA of interest (12). The desired product DNA can be isolated by selecting for loss of the selectable marker genes (9) and (10)

located between respectively the first (4) and second (5) recombination sites and the third (6) and fourth (7) recombination sites. Optionally, an additional selectable marker gene may be included between the second (5) and third (6) recombination site to allow selection for the presence of this selectable marker gene and consequently for the optional intron sequence, which is flanked by functional intron processing signal sequences (11). The acceptor vector, as well as the product vector further comprises an origin of replication (Ori; (1)) and a selectable marker gene (2) to allow selection for the presence of the plasmid.

[036] This results in a chimeric DNA construct with the desired configuration comprising a eukaryotic promoter region (3); a recombination site (15) produced by the recombination between recombination sites (4) and (13); a first copy of the DNA of interest (12); a recombination site (16) produced by the recombination between recombination sites (5) and (14); optionally an intron sequence flanked by intron processing signals (11); a recombination site (17) produced by the recombination between recombination sites (6) and (14); a second copy of the DNA of interest (12) in opposite orientation to the first copy of the DNA of interest; a recombination site (18) produced by the recombination between recombination sites (7) and (13); a eukaryotic transcription terminator and polyadenylation signal (8).

[037] Fig 2A: A nucleic acid of interest (12) is amplified by PCR using primers comprising two different recombination sites which upon recombination with the recombination sites on an intermediate vector (Fig 2B) will yield recombination sites compatible with the first (4) and fourth (5) and with the second (6) and third (7) recombination site on the acceptor vector respectively.

[038] Fig 2B: The insert DNA obtained in Fig 2A is allowed to recombine with the intermediate vector in the presence of at least one recombination protein to obtain an intermediate DNA wherein the DNA of interest (12) is flanked by two different recombination sites (13, 14) and which further comprises an origin of replication (1) and a selectable marker gene (2).

[039] Fig 2C: The intermediate DNA is then allowed to recombine with the acceptor vector using at least one second recombination protein (basically as described for Fig 1B).

[040] Fig 3: Schematic representation of the acceptor vector "pHELLSGATE"

[041] Fig 4: Schematic representation of the acceptor vectors "pHELLSGATE 8" "pHELLSGATE 11" and "pHELLSGATE 12".

[042] DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

[043] The current invention is based on the unexpected finding by the inventors that recombinational cloning was an efficient one-step method to convert a nucleic acid fragment of interest into a chimeric DNA construct capable of producing a dsRNA transcript comprising a sense and antisense nucleotide sequence capable of being expressed in eukaryotic cells. The dsRNA molecules are efficient effectors of gene silencing. These methods improve the efficiency problems previously encountered to produce chimeric DNAs with long inverted repeats.

[044] Thus, in a first embodiment, the invention provides a method for making a chimeric DNA construct or chimeric gene capable of expressing an RNA transcript in a eukaryotic cell , the RNA being capable of internal basepairing between a stretch of nucleotides corresponding to a nucleic acid of interest and its complement (i.e. the stretch of nucleotides in inverted orientation) located elsewhere in the transcript (and thus forming a hairpin RNA) comprising the following steps:

providing an "acceptor vector" comprising the following operably linked DNA fragments:

an origin of replication allowing replication in a host cell (1),

a selectable marker region (2) capable of being expressed in the host cell; and

a chimeric DNA construct comprising in sequence:

a promoter or promoter region (3) capable of being recognized by RNA polymerases of a eukaryotic cell;

a first recombination site (4), a second recombination site (5), a third recombination site (6) and a fourth recombination site (7), whereby

the first (4) and fourth recombination site (7) are capable of reacting with the same other recombination site and preferably are identical to each other;

the second (5) and third (6) recombination site are also capable of reacting with the same other recombination site and preferably are identical to each other;

the first (4) and second (5) recombination site do not recombine with each other or with the same other recombination site; and

the third (6) and fourth (7) recombination site do not recombine with each other or with the same other recombination site; and

a 3' transcription terminating and polyadenylation region (8) functional in a eukaryotic cell;

providing an "insert DNA" comprising

the DNA segment of interest (12) flanked by

a fifth recombination site (13) which is capable of recombining with the first(4) or fourth (7) recombination site but preferably not with the second (5) or third (6) recombination site;

a sixth recombination site (14), which is capable of recombining with the second (5) or third (6) recombination site but preferably not with the first (4) or fourth (7) recombination site.

combining *in vitro* the insert DNA and the acceptor vector in the presence of at least one specific recombination protein; and

allowing the recombination to occur to produce a reaction mixture comprising *inter alia* "product DNA" molecules which comprise in sequence

the promoter or promoter region (3) capable of being recognized by RNA polymerases of a eukaryotic cell;

a recombination site (15) which is the recombination product of the first (4) and fifth recombination site (13);

a first copy of the DNA fragment of interest (12);

a recombination site (16) which is the recombination product of the second (4) and the sixth recombination site (14);

a recombination site (17) which is the recombination product of the third (5) and the sixth recombination site (14);

a second copy of the DNA fragment of interest in opposite orientation (12) with regard to the first copy ;

a recombination site (18) which is the recombination product of the fourth (7) and the fifth recombination site (13); and

a 3' transcription terminating and polyadenylation region (8) functional in a eukaryotic cell; and

selecting the product DNA molecules.

This method is schematically outlined in Figure 1, with non-limiting examples of recombination sites and selectable markers.

[045] As used herein, a "host cell" is any prokaryotic or eukaryotic organism that can be a recipient for the acceptor vector or the product DNA. Conveniently, the host cell will be an *Escherichia coli* strain commonly used in recombinant DNA methods.

[046] A "recombination protein" is used herein to collectively refer to site-specific recombinases and associated proteins and/or co-factors. Site-specific recombinases are enzymes that are present in some viruses and bacteria and have been characterized to have both endonuclease and ligase properties. These recombinases (along with associated proteins in some cases) recognize specific sequences of bases in DNA and exchange the DNA segments flanking those segments. Various recombination proteins are described in the art(see WO 96/40724 herein incorporated by reference in its entirety, at least on page 22 to 26). Examples of such recombinases include Cre from bacteriophage P1 and Integrase from bacteriophage lambda.

[047] Cre is a protein from bacteriophage P1 (Abremski and Hoess, 1984) which catalyzes the exchange between 34 bp DNA sequences called *lox*P sites (see Hoess et al., 1986. Cre is available commercially (Novagen, Catalog 69247-1).

[048] Integrase (Int) is a protein from bacteriophage lambda that mediates the integration of the lambda genome into the *E. coli* chromosome. The bacteriophage lambda Int recombinational proteins promote irreversible recombination between its substrate *att* sites as part of the formation or induction of a lysogenic state. Reversibility of the recombination reactions results from two independent pathways for integrative or excisive recombination. Cooperative and competitive interactions involving four proteins (Int, Xis, IHF and FIS) determine the direction of recombination. Integrative recombination involves the Int and IHF proteins and *attP* (240bp) and *att*B (25b) recombination sites. Recombination results in the formation of two new sites: *att*L and *att*R. A commercial preparation comprising Int and IHF proteins is commercially available (BP clonase™ ; Life Technologies). Excisive recombination requires Int, IHF, and Xis and sites *att*L and *att*R to generate *att*P and *att*B. A commercial preparation comprising Int, IHF and Xis proteins is commercially available (LR clonase™ ; Life Technologies).

[049] A "recombination site" as used herein refers to particular DNA sequences, which a recombinase and possibly associated proteins recognizes and binds. The recombination site recognized by Cre recombinase is *lox*P which is a 34 base pair sequence comprised of two 13 base pair inverted repeats (serving as recombinase binding sites) flanking an 8 base pair core sequence. The recombination sites *att*B, *att*P, *att*L and *att*R are recognized by lambda integrase. *Att*B is an approximately 25 base pair sequence containing two 9 base pair core-type Int binding sites and a 7 base pair overlap region. *Att*P is an approximately 240 base pair sequence containing core-type Int binding sites and arm-type Int binding sites as well as sites for auxiliary proteins IHF, FIS and Xis (Landy 1993). Each of the *att* sites contains a 15 bp core sequence with individual sequence elements of functional significance lying within, outside and across the boundaries of this common core (Landy, 1989) Efficient recombination between the various att sites requires that the sequence of the central common region is substantially

identical between the recombining partners. The exact sequence however is modifiable as disclosed in WO 96/40724 and the variant recombination sites selected from

*att*B1: AGCCTGCTTTTTTGTACAAACTTGT (SEQ ID No 1);

*att*B2: AGCCTGCTTTCTTGTACAAACTTGT (SEQ ID No 2);

*att*B3: ACCCAGCTTTCTTGTACAAACTTGT (SEQ ID No 3);

*att*R1: GTTCAGCTTTTTTGTACAAACTTGT (SEQ ID No 4);

*att*R2: GTTCAGCTTTCTTGTACAAACTTGT (SEQ ID No 5);

*att*R3: GTTCAGCTTTCTTGTACAAAGTTGG (SEQ ID No 6);

*att*L1: AGCCTGCTTTTTTGTACAAAGTTGG (SEQ ID No 7);

*att*L2: AGCCTGCTTTCTTGTACAAAGTTGG (SEQ ID No 8);

*att*L3: ACCCAGCTTTCTTGTACAAAGTTGG (SEQ ID No 9);

*att*P1: GTTCAGCTTTTTTGTACAAAGTTGG (SEQ ID No 10); or

*att*P2,P3: GTTCAGCTTTCTTGTACAAAGTTGG (SEQ ID No 11)

allow more flexibility in the choice of suitable pairs or recombination sites that have the capability to recombine (as indicated by their index number).

[050] It will be clear to the skilled artisan that a correspondence is required between the recombination site(s) used and the recombination proteins used.

[051] In one embodiment, the following combinations of recombination sites for the acceptor vector are present in the acceptor vector:

the first (4) and fourth (7) recombination sites are identical and comprise

*att*P1 comprising

the nucleotide sequence of SEQ ID No 10 and

the second (5) and third (6) recombination site are also identical and comprise

*att*P2 comprising

the nucleotide sequence of SEQ ID No 11; or

the first (4) and fourth (7) recombination sites are identical and comprise

*att*R1 comprising

the nucleotide sequence of SEQ ID No 4 and

the second (5) and third (6) recombination site are also identical and comprise

*att*R2 comprising

the nucleotide sequence of SEQ ID No 5; and

the following combinations of recombination sites for the insert DNA are used:

the fifth (13) recombination site comprises

*att*B1 comprising

the nucleotide sequence of SEQ ID No 1 and

the sixth (14) recombination site comprises

*att*B2 comprising

the nucleotide sequence of SEQ ID No 2,

the combination being suitable for recombination with the first acceptor vector mentioned above; or

the fifth (13) recombination site comprises

*att*L1 comprising

the nucleotide sequence of SEQ ID No 7 and

the sixth (14) recombination site comprises

*att*L2 comprising

the nucleotide sequence of SEQ ID No 8,

the combination being suitable for recombination with the second acceptor vector mentioned above.

[052] It has been unexpectedly found that product DNA molecules (resulting from recombination between the above mentioned second acceptor vector with attR recombination sites (such as pHELLSGATE 8) and insert DNA flanked by attL recombination sites) wherein the gene inserts in both orientations are flanked by attB recombination sites are more effective in

silencing of the target gene(both quantitatively and qualitatively) than product DNA molecules (resulting from recombination between the above mentioned first acceptor vector with attP recombination sites (such as pHELLSGATE or pHELLSGATE 4) and insert DNA flanked by attB recombination sites) wherein the gene inserts in both orientations are flanked by attL recombination sites. Although not intending to limit the invention to a particular mode of action it is thought that the greater length of the attL sites and potential secondary structures therein may act to inhibit transcription yielding the required dsRNA to a certain extent. However, acceptor vectors such as the above mentioned first acceptor vectors with attP sites may be used when target gene silencing to a lesser extent would be useful or required.

[053] The dsRNA obtained by the chimeric DNA construct made according to the invention may be used to silence a nucleic acid of interest, *i.e.,* to reduce its phenotypic expression, in a eukaryotic organism, particularly a plant, either directly or by transcription of the chimeric DNA construct in the cells of the eukaryotic organism. When this is the case, the following considerations may apply.

[054] The length of the nucleic acid of interest (12) may vary from about 10 nucleotides (nt) up to a length equaling the length (in nucleotides) of the target nucleic acid whose phenotypic expression is to be reduced. Preferably the total length of the sense nucleotide sequence is at least 10 nt, or at least 19 nt, or at least 21 nt, or at least 25 nt, or at least about 50 nt, or at least about 100 nt, or at least about 150 nt, or at least about 200 nt, or at least about 500 nt. It is expected that there is no upper limit to the total length of the sense nucleotide sequence, other than the total length of the target nucleic acid. However for practical reasons (such as, *e.g.,* stability of the chimeric genes) it is expected that the length of the sense nucleotide sequence should not exceed 5000 nt, particularly should not exceed 2500 nt and could be limited to about 1000 nt.

[055] It will be appreciated that the longer the total length of the nucleic acid of interest (12), the less stringent the requirements for sequence identity between the nucleic acid of interest

and the corresponding sequence in the target gene. Preferably, the nucleic acid of interest should have a sequence identity of at least about 75% with the corresponding target sequence, particularly at least about 80 %, more particularly at least about 85%, quite particularly about 90%, especially about 95%, more especially about 100%, quite especially be identical to the corresponding part of the target nucleic acid. However, it is preferred that the nucleic acid of interest always includes a sequence of about 10 consecutive nucleotides, particularly about 25 nt, more particularly about 50 nt, especially about 100 nt, quite especially about 150 nt with 100% sequence identity to the corresponding part of the target nucleic acid. Preferably, for calculating the sequence identity and designing the corresponding sense sequence, the number of gaps should be minimized, particularly for the shorter sense sequences.

[056] For the purpose of this invention, the "sequence identity" of two related nucleotide or amino acid sequences, expressed as a percentage, refers to the number of positions in the two optimally aligned sequences which have identical residues (x100) divided by the number of positions compared. A gap, *i.e.* a position in an alignment where a residue is present in one sequence but not in the other is regarded as a position with non-identical residues. The alignment of the two sequences is performed by the Needleman and Wunsch algorithm (Needleman and Wunsch 1970). The computer-assisted sequence alignment above, can be conveniently performed using standard software program such as GAP, which is part of the Wisconsin Package Version 10.1 (Genetics Computer Group, Madison, Wisconsin, USA) using the default scoring matrix with a gap creation penalty of 50 and a gap extension penalty of 3. Sequences are indicated as “essentially similar” when such sequence have a sequence identity of at least about 75%, particularly at least about 80 %, more particularly at least about 85%, quite particularly about 90%, especially about 95%, more especially about 100%, quite especially are identical. It is clear than when RNA sequences are the to be essentially similar or have a certain degree of sequence identity with DNA sequences, thymine (T) in the DNA sequence is considered equal to uracil (U) in the RNA sequence.

[057] The "insert DNA" may conveniently be provided using DNA amplification procedures, such as PCR, of the nucleic acid of interest, using as primers oligonucleotide sequences incorporating appropriate recombination sites as well as oligonucleotide sequences appropriate for the amplification of the nucleic acid of interest. However, alternative methods are available in the art to provide the nucleic acid of interest with the flanking recombination sites, including but not limited to covalently linking oligonucleotides or nucleic acid fragments comprising such recombination sites to the nucleic acid(s) of interest using ligase(s).

[058] The providing of the appropriate flanking recombination sites to the nucleic acid may also proceed in several steps. For example, in a first step the flanking sites provided to the nucleic acid of interest may be such that upon recombination with the recombination sites in an intermediate vector new recombination sites are created flanking the nucleic acid of interest, now compatible for recombination with the acceptor vector. This scheme is outlined in Fig 2, with non-limiting examples of recombination sites and selectable markers. It is understood that the insert DNA may be in a circular form or in a linear form.

[059] As used herein, an "origin of replication" is a DNA fragment which allows replication of the acceptor vector in microorganisms, preferably bacteria, particularly *E. coli* strains, and ensures that upon multiplication of the microorganism, the daughter cells receive copies of the acceptor vector.

[060] "Selectable marker (gene)" is used herein to indicate a DNA segment that allows selection or screening for the presence or absence of that DNA segment under suitable conditions. Selectable markers include but are not limited to:

DNA segments that encode products which provide resistance against otherwise toxic compounds (*e.g.* antibiotic resistance genes, herbicide resistance genes);

DNA segments encoding products which are otherwise lacking in the recipient cell (e.g. tRNA genes, auxotrophic markers);

DNA segments encoding products which suppress the activity of a gene product;

DNA segments encoding products which can readily be identified (*e.g.* β-galactosidase, green fluorescent protein (GFP), β-glucuronidase (GUS));

DNA segments that bind products which are otherwise detrimental to cell survival and/or function;

DNA segments that are capable of inhibiting the activity of any of the DNA segments Nos (1) to (5) (*e.g.* antisense oligonucleotides);

DNA segments that bind products that modify a substrate (*e.g.* restriction endonuclease);

DNA segments that can be used to isolate a desired molecule (e.g. specific protein binding sites);

DNA segments that encode a specific nucleotide sequence which can be otherwise non-functional (*e.g.* for PCR amplification of subpopulations of molecules);

DNA segments, which when absent, directly or indirectly confer sensitivity to particular compound(s); and/or

DNA segments, which when absent, directly or indirectly confer resistance to particular compound(s).

[061] Preferred first selectable markers (2) are antibiotic resistance genes. A large number of antibiotic resistance genes, particularly which can be used in bacteria, are available in the art and include but are not limited to aminoglycoside phosphotransferase I and II, chloramphenicol acetyltransferase, beta-lactamase, and/or aminoglycoside adenosyltransferase.

[062] Preferred second selectable markers (9) and third selectable markers (10) are selectable markers allowing a positive selection when absent or deleted after recombination (*i.e.* in the product DNA) such as but not limited to *ccd*B gene the product of which interferes with *E. coli* DNA gyrase and thereby inhibits growth of most *E. coli* strains. Preferably, the second and third markers are identical.

[063] In one embodiment of the invention, the acceptor comprises a fourth selectable marker (19) between the second (5) and third (6) recombination site, preferably a marker allowing positive selection for the presence thereof, such as a antibiotic resistance gene, *e.g.* chloramphenicol resistance gene. Preferably, the fourth selectable marker should be different from first selectable marker and different from the second and third selectable marker. The presence of a fourth selectable marker allows to select or screen for the retention of the DNA region between the second (5) and third (6) recombination site in the product DNA. This increases the efficiency with which the desired product DNAs having the nucleic acid of interest cloned in inverted repeat and operably linked to eukaryotic expression signals may be obtained. However, it has been found that with most of the acceptor vectors tested, the presence of a selectable marker is not required and has little influence on the ratio of expected and desired product DNA molecules (which usually exceeds about 90% of obtained product DNA molecules) to undesired product DNA molecules.

[064] It will be understood that a person skilled in the art has a number of techniques available for recognizing the expected and desired product DNA molecules, such as but not limited to restriction enzyme digests or even determining the nucleotide sequence of the recombination product.

[065] In another embodiment of the invention, the acceptor vector further comprises a pair of intron processing signals (11) or an intron sequence functional in the eukaryotic cell, preferably located between the second (5) and third (6) recombination site. However, the pair of intron processing signals or the intron may also be located elsewhere in the chimeric construct between the promoter or promoter region (3) and the terminator region (8). As indicated in the background art, this will improve the efficiency with which the chimeric DNA construct encoding the dsRNA will be capable of reducing the phenotypic expression of the target gene in the eukaryotic cell. A particularly preferred intron functional in cells of plants is the *pdk* intron (*Flaveria trinervia* pyruvate orthophosphate dikinase intron 2 ; see WO99/53050 incorporated by

reference). The fourth selectable marker (19) may be located between the intron processing signals or within the intron (if these are located between the second and third recombination site), but may also be located adjacent to the intron processing signals or the intron.

[066] A person skilled in the art will recognize that the product DNA molecules, resulting from a recombination with an acceptor vector as herein described, which comprise a region between the second (5) and third (6) recombination will fall into two classes which can be recognized by virtue of the orientation of that intervening region. In the embodiments wherein the acceptor vector also comprises an intron, the different orientation may necessitate an additional step of identifying the correct orientation. To avoid this additional step, the acceptor vector may comprise an intron that can be spliced out independent of its orientation (such as present in pHELLSGATE 11) or the acceptor vector may comprise a spliceable intron in both orientations (such as present in pHELLSGATE 12).

[067] As used herein, the term "promoter" denotes any DNA that is recognized and bound (directly or indirectly) by a DNA-dependent RNA-polymerase during initiation of transcription. A promoter includes the transcription initiation site, and binding sites for transcription initiation factors and RNA polymerase, and can comprise various other sites (*e.g.,* enhancers), at which gene expression regulatory proteins may bind.

[068] The term "regulatory region", as used herein, means any DNA that is involved in driving transcription and controlling (*i.e.,* regulating) the timing and level of transcription of a given DNA sequence, such as a DNA coding for a protein or polypeptide. For example, a 5' regulatory region (or "promoter region") is a DNA sequence located upstream (*i.e.,* 5') of a coding sequence and which comprises the promoter and the 5'-untranslated leader sequence. A 3' regulatory region is a DNA sequence located downstream (*i.e.,* 3') of the coding sequence and which comprises suitable transcription termination (and/or regulation) signals, including one or more polyadenylation signals.

[069] As used herein, the term "plant-expressible promoter" means a DNA sequence that is capable of controlling (initiating) transcription in a plant cell. This includes any promoter of plant origin, but also any promoter of non-plant origin which is capable of directing transcription in a plant cell, i.e., certain promoters of viral or bacterial origin such as the CaMV35S, the subterranean clover virus promoter No 4 or No 7, or T-DNA gene promoters. Other suitable promoters include tissue-specific or organ-specific promoters including but not limited to seed-specific promoters (e.g., WO89/03887), organ-primordia specific promoters (An et al., 1996), stem-specific promoters (Keller et al., 1988), leaf specific promoters (Hudspeth et al., 1989), mesophyl-specific promoters (such as the light-inducible Rubisco promoters), root-specific promoters (Keller et al.,1989), tuber-specific promoters (Keil et al., 1989), vascular tissue specific promoters (Peleman et al., 1989), stamen-selective promoters (WO 89/10396, WO 92/13956), dehiscence zone specific promoters (WO 97/13865) and the like.

[070] The acceptor vector may further comprise a selectable marker for expression in a eukaryotic cell. Selectable marker genes for expression in eukaryotic cells are well known in the art, including but not limited to chimeric marker genes. The chimeric marker gene can comprise a marker DNA that is operably linked at its 5' end to a promoter, functioning in the host cell of interest, particularly a plant-expressible promoter, preferably a constitutive promoter, such as the CaMV 35S promoter, or a light inducible promoter such as the promoter of the gene encoding the small subunit of Rubisco; and operably linked at its 3' end to suitable plant transcription 3' end formation and polyadenylation signals. It is expected that the choice of the marker DNA is not critical, and any suitable marker DNA can be used. For example, a marker DNA can encode a protein that provides a distinguishable color to the transformed plant cell, such as the A1 gene (Meyer *et al.*, 1987), can provide herbicide resistance to the transformed plant cell, such as the *bar* gene, encoding resistance to phosphinothricin (EP 0,242,246), or can provide antibiotic resistance to the transformed cells, such as the *aac(6')* gene, encoding resistance to gentamycin (WO94/01560).

[071] The acceptor vector may also further comprise left and right T-DNA border sequences flanking the chimeric DNA construct, and may comprise an origin of replication functional in *Agrobacterium* spp. and/or a DNA region of homology with a helper Ti-plasmid as described in EP 0 116 718.

[072] The efficiency and ease by which any nucleic acid of interest may be converted into a chimeric DNA construct comprising two copies of the nucleic acid of interest in inverted repeat and operably linked to eukaryotic 5' and 3' regulatory regions using the means and methods according to the invention, makes these particularly apt for automation and high throughput analysis.

[073] It will be clear to the person skilled in the art that the acceptor vectors as hereinbefore described can be readily adapted to provide a vector which can be used to produce *in vitro* large amounts of double stranded RNA or RNAi comprising a complementary sense and antisense portion essentially similar to a target gene of choice as described elsewhere in this application, by exchanging the promoter capable of being expressed in a eukaryotic cell for a promoter recognized by any RNA polymerase. Very suitable promoters to this end are the promoters recognized by bacteriophage single subunit RNA polymerases such as the promoters recognized by bacteriophage single subunit RNA polymerase such as the RNA polymerases derived from the *E. coli* phages T7, T3, φI, φII, W31, H, Y, A1, 122, cro, C21, C22, and C2; *Pseudomonas putida* phage gh-1; *Salmonella typhimurium* phage SP6; *Serratia marcescens* phage IV; *Citrobacter* phage VIIII; and *Klebsiella* phage No.11 (Hausmann, *Current Topics in Microbiology and Immunology,* 75: 77-109 (1976); Korsten et al., *J. Gen Virol.* 43: 57-73 (1975); Dunn et al., *Nature New Biology,* 230: 94-96 (1971); Towle et al., *J. Biol. Chem.* 250: 1723-1733 (1975); Butler and Chamberlin, *J. Biol. Chem.,* 257: 5772-5778 (1982)). Examples of such promoters are a T3 RNA polymerase specific promoter and a T7 RNA polymerase specific promoter, respectively. A T3 promoter to be used as a first promoter in the CIG can be any promoter of the T3 genes as described by McGraw et al, *Nucl. Acid Res.* 13: 6753-6766 (1985).

Alternatively, a T3 promoter may be a T7 promoter that is modified at nucleotide positions -10, -11 and -12 in order to be recognized by T3 RNA polymerase (Klement et al., *J. Mol. Biol.* 215, 21-29 (1990)). A preferred T3 promoter is the promoter having the "consensus" sequence for a T3 promoter, as described in US Patent 5,037,745. A T7 promoter which may be used according to the invention, in combination with T7 RNA polymerase, comprises a promoter of one of the T7 genes as described by Dunn and Studier, *J. Mol. Biol.* 166: 477-535 (1983). A preferred T7 promoter is the promoter having the "consensus" sequence for a T7 promoter, as described by Dunn and Studier (*supra*).

[074] Thus, the invention also provides an acceptor vector comprising:

origin of replication allowing replication in a host cell (1);

a selectable marker region (2) capable of being expressed in the host cell; and

a chimeric DNA construct comprising in sequence:

a promoter or promoter region (3) capable of being recognized by a bacteriophage single subunit RNA polymerase;

a first recombination site (4), a second recombination site (5), a third recombination site (6) and a fourth recombination site (7) whereby

the first (4) and fourth recombination site (7) are capable of reacting with the same other recombination site and preferably are identical to each other;

the second (5) and third (6) recombination site are also capable of reacting with the same other recombination site and preferably are identical to each other;

the first (4) and second (5) recombination site do not recombine with each other or with the same other recombination site; and

the third (6) and fourth (7) recombination site do not recombine with each other or with the same other recombination site; and

a 3' transcription terminating and polyadenylation region (8) functional in a eukaryotic cell.

[075] The acceptor vector may be used to convert a DNA fragment of interest into an inverted repeat structure as described elsewhere in the application and dsRNA can be produced in large amounts by contacting the acceptor vector DNA with the appropriate bacteriophage single subunit RNA polymerase under conditions well known to the skilled artisan. The so-produced dsRNA can then be used for delivery into cells prone to gene silencing, such as plant cells, fungal cells or animal cells. dsRNA may be introduced in animal cells via liposomes or other transfection agents (*e.g.* Clonfection transfection reagent or the CalPhos Mammalian transfection kit from ClonTech) and could be used for methods of treatment of animals, including humans, by silencing the appropriate target genes.

[076] The acceptor vectors may also be equipped with any prokaryotic promoter suitable for expression of dsRNA in a particular prokaryotic host. The prokaryotic host can be used as a source of dsRNA, *e.g.* by feeding it to an animal, such as a nematode, in which the silencing of the target gene is envisioned.

[077] The promoter capable of expression in eukaryotic cell may also be a promoter capable of expression in a mammalian cell and vectors according to the invention may transiently be delivered using a retroviral delivery system or other animal transfection system.

[078] In another embodiment of the invention, a method is provided for making a eukaryotic organism, particularly a plant, wherein the phenotypic expression of a target nucleic acid of interest is reduced or inhibited, comprising the steps of preparing a chimeric DNA construct comprising a nucleic acid of interest (12) comprising a nucleotide sequence of at least 19 bp or 25 bp having at least 70% sequence identity to the target nucleic acid of interest and capable of expressing a dsRNA in cells of the eukaryotic organism, particularly a plant according to the methods of the current invention and introducing the chimeric DNA construct in cells of the eukaryotic organism, and isolating eukaryotic organism transgenic for the chimeric DNA construct.

[079] As used herein, "phenotypic expression of a target nucleic acid of interest" refers to any quantitative trait associated with the molecular expression of a nucleic acid in a host cell and may thus include the quantity of RNA molecules transcribed or replicated, the quantity of post-transcriptionally modified RNA molecules, the quantity of translated peptides or proteins, the activity of such peptides or proteins.

[080] A "phenotypic trait" associated with the phenotypic expression of a nucleic acid of interest refers to any quantitative or qualitative trait, including the trait mentioned, as well as the direct or indirect effect mediated upon the cell, or the organism containing that cell, by the presence of the RNA molecules, peptide or protein, or posttranslationally modified peptide or protein. The mere presence of a nucleic acid in a host cell, is not considered a phenotypic expression or a phenotypic trait of that nucleic acid, even though it can be quantitatively or qualitatively traced. Examples of direct or indirect effects mediated on cells or organisms are, *e.g.,* agronomically or industrial useful traits, such as resistance to a pest or disease; higher or modified oil content etc.

[081] As used herein, "reduction of phenotypic expression" refers to the comparison of the phenotypic expression of the target nucleic acid of interest to the eukaryotic cell in the presence of the RNA or chimeric genes of the invention, to the phenotypic expression of the target nucleic acid of interest in the absence of the RNA or chimeric genes of the invention. The phenotypic expression in the presence of the chimeric RNA of the invention should thus be lower than the phenotypic expression in absence thereof, preferably be only about 25%, particularly only about 10%, more particularly only about 5% of the phenotypic expression in absence of the chimeric RNA, especially the phenotypic expression should be completely inhibited for all practical purposes by the presence of the chimeric RNA or the chimeric gene encoding such an RNA.

[082] A reduction of phenotypic expression of a nucleic acid where the phenotype is a qualitative trait means that in the presence of the chimeric RNA or gene of the invention, the phenotypic trait switches to a different discrete state when compared to a situation in which

such RNA or gene is absent. A reduction of phenotypic expression of a nucleic acid may thus, *i.a.* be measured as a reduction in transcription of (part of) that nucleic acid, a reduction in translation of (part of) that nucleic acid or a reduction in the effect the presence of the transcribed RNA(s) or translated polypeptide(s) have on the eukaryotic cell or the organism, and will ultimately lead to altered phenotypic traits. It is clear that the reduction in phenotypic expression of a target nucleic acid of interest, may be accompanied by or correlated to an increase in a phenotypic trait.

[083] As used herein a "target nucleic acid of interest" refers to any particular RNA molecule or DNA sequence which may be present in a eukaryotic cell, particularly a plant cell whether it is an endogenous nucleic acid, a transgenic nucleic acid, a viral nucleic acid, or the like.

[084] Methods for making transgenic eukaryotic organisms, particularly plants are well known in the art. Gene transfer can be carried out with a vector that is a disarmed Ti-plasmid, comprising a chimeric gene of the invention, and carried by *Agrobacterium.* This transformation can be carried out using the procedures described, for example, in EP 0 116 718. Particular kinds of Agrobacterium mediated transformation methods are the so-called *in planta* methods, which are particularly suited for *Arabidopsis* spp. transformation (*e.g.,* Clough and Bent, *Plant J.*16:735-534, 1998). Alternatively, any type of vector can be used to transform the plant cell, applying methods such as direct gene transfer (as described, for example, in EP 0 233 247), pollen-mediated transformation (as described, for example, in EP 0 270 356, WO85/01856 and US 4,684,611), plant RNA virus-mediated transformation (as described, for example, in EP 0 067 553 and US 4,407,956), liposome-mediated transformation (as described, for example, in US 4,536,475), and the like. Other methods, such as microprojectile bombardment, as described for corn by Fromm *et al.* (1990) and Gordon-Kamm *et al.* (1990), are suitable as well. Cells of monocotyledonous plants, such as the major cereals, can also be transformed using wounded and/or enzyme-degraded compact embryogenic tissue capable of forming compact embryogenic callus, or wounded and/or degraded immature embryos as described in

WO92/09696. The resulting transformed plant cell can then be used to regenerate a transformed plant in a conventional manner.

[085] The obtained transformed plant can be used in a conventional breeding scheme to produce more transformed plants with the same characteristics or to introduce the chimeric gene for reduction of the phenotypic expression of a nucleic acid of interest of the invention in other varieties of the same or related plant species, or in hybrid plants. Seeds obtained from the transformed plants contain the chimeric genes of the invention as a stable genomic insert.

[086] In another embodiment, the invention provides a method for isolating a nucleic acid molecule involved in determining a particular phenotypic trait of interest. The method involves the following steps:

preparing a library of chimeric DNA constructs capable of expressing a dsRNA in cells of the eukaryotic non-human organism using the methods and means described in the current invention;

introducing individual representatives of this library of chimeric DNA constructs in cells of the eukaryotic non-human organism, preferably by stable integration in their genome, particularly their nuclear genome;

isolating a eukaryotic organism exhibiting the particular trait; and

isolating the corresponding nucleic acid molecule present in the eukaryotic organism with the trait of interest, preferably from the aforementioned library.

[087] It will be understood that the methods and means of the invention may be used to determine the function of an isolated nucleic acid fragment or sequence with unknown function, by converting a part or the whole of that nucleic acid fragment or sequence according to the methods of the invention into a chimeric construct capable of making a dsRNA transcript when introduced in a eukaryotic cell, introducing that chimeric DNA construct into a eukaryotic organism to isolate preferably a number of transgenic organisms and observing changes in phenotypic traits.

[088] The invention also provides acceptor vectors, as described in this specification as well as kits comprising such vectors.

[089] It will be understood that the vectors, methods and kits according to the invention may be used in all eukaryotic organisms which are prone to gene silencing including yeast, fungi, plants, animals such as nematodes, insects and arthropods, vertebrates including mammals and humans.

[090] Also provided by the invention are non-human organisms comprising chimeric DNA constructs comprising in sequence the following operably linked DNA fragments a promoter or promoter region (3) capable of being recognized by RNA polymerases of the eukaryotic cell; a recombination site (15) which is the recombination product of the first (4) recombination site on the acceptor vector and the fifth recombination site (13) flanking the DNA of interest; a first DNA copy of the nucleic acid fragment of interest (12);

a recombination site (16) which is the recombination product of the second (4) recombination site on the acceptor vector and the sixth recombination site (14) flanking the DNA of interest;

a recombination site (17) which is the recombination product of the third (5) recombination site on the acceptor vector and the sixth recombination site (14) flanking the DNA of interest;

a second DNA copy of the nucleic acid fragment of interest in opposite orientation (12) compared to the first copy;

a recombination site (18) which is the recombination product of the fourth (7) recombination site on the acceptor vector and the fifth recombination site (13) flanking the DNA of interest; and

a 3' transcription terminating and polyadenylation region (8) functional in a eukaryotic cell.

[091] As used herein "comprising" is to be interpreted as specifying the presence of the stated features, integers, steps or components as referred to, but does not preclude the presence or addition of one or more features, integers, steps or components, or groups thereof. Thus, *e.g.,* a nucleic acid or protein comprising a sequence of nucleotides or amino acids, may comprise more nucleotides or amino acids than the actually cited ones, *i.e.,* be embedded in a larger nucleic acid or protein. A chimeric gene comprising a DNA region that is functionally or structurally defined, may comprise additional DNA regions etc.

[092] The term "gene" means any DNA fragment comprising a DNA region (the "transcribed DNA region") that is transcribed into a RNA molecule (*e.g.,* a mRNA) in a cell operably linked to suitable regulatory regions, *e.g.,* a plant-expressible promoter. A gene may thus comprise several operably linked DNA fragments such as a promoter, a 5' leader sequence, a coding region, and a 3' region comprising a polyadenylation site. A plant gene endogenous to a particular plant species (endogenous plant gene) is a gene which is naturally found in that plant species or which can be introduced in that plant species by conventional breeding. A chimeric gene is any gene that is not normally found in a plant species or, alternatively, any gene in which the promoter is not associated in nature with part or all of the transcribed DNA region or with at least one other regulatory region of the gene.

[093] The term "expression of a gene" refers to the process wherein a DNA region which is operably linked to appropriate regulatory regions, particularly to a promoter, is transcribed into an RNA which is biologically active *i.e.,* which is either capable of interaction with another nucleic acid or which is capable of being translated into a polypeptide or protein. A gene is the to encode an RNA when the end product of the expression of the gene is biologically active RNA, such as *e.g.* an antisense RNA, a ribozyme or a replicative intermediate. A gene is the to encode a protein when the product of the expression of the gene is a protein or polypeptide.

[094] A nucleic acid is "capable of being expressed", when the nucleic acid, when introduced in a suitable host cell, particularly in a plant cell, can be transcribed (or replicated) to yield an RNA, and/or translated to yield a polypeptide or protein in that host cell.

[095] The following non-limiting Examples describe the construction of acceptor vectors and the application thereof for the conversion of nucleic acid fragments of interest into chimeric DNA constructs capable of expressing a dsRNA transcript in eukaryotic cells. Unless stated otherwise in the Examples, all recombinant DNA techniques are carried out according to standard protocols as described in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, NY and in Volumes 1 and 2 of Ausubel *et al.* (1994) *Current Protocols in Molecular Biology, Current Protocols*, USA. Standard materials and methods for plant molecular work are described in *Plant Molecular Biology Labfax* (1993) by R.D.D. Croy, jointly published by BIOS Scientific Publications Ltd (UK) and Blackwell Scientific Publications, UK. Other references for standard molecular biology techniques include Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual,* Third Edition, Cold Spring Harbor Laboratory Press, NY, Volumes I and II of Brown (1998) *Molecular Biology LabFax*, Second Edition, Academic Press (UK). Standard materials and methods for polymerase chain reactions can be found in Dieffenbach and Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, and in McPherson at al. (2000) *PCR - Basics: From Background to Bench,* First Edition, Springer Verlag, Germany.

[096] Throughout the description and Examples, reference is made to the following sequences:

SEQ ID N° 1: core sequence of recombination site *att*B1

SEQ ID No 2: core sequence of recombination site *att*B2

SEQ ID No 3: core sequence of recombination site *att*B3

SEQ ID No 4: core sequence of recombination site *att*R1

SEQ ID No 5: core sequence of recombination site *att*R2

SEQ ID No 6: core sequence of recombination site *att*R3

SEQ ID No 7: core sequence of recombination site *att*L1

SEQ ID No 8: core sequence of recombination site *att*L2

SEQ ID No 9: core sequence of recombination site *att*L3

SEQ ID No 10: core sequence of recombination site *att*P1

SEQ ID No 11: core sequence of recombination sites *att*P2,P3

SEQ ID No 12: nucleotide sequence of chalcone synthase gene of Arabidopsis

SEQ ID No 13: nucleotide sequence of the acceptor vector "pHELLSGATE"

SEQ ID No 14: oligonucleotide attB1 "forward" primer used for amplification of 400bp and 200 bp CHS fragments.

SEQ ID No 15: oligonucleotide attB2 "reverse" primer for amplification of the 400 bp CHS fragment.

SEQ ID No 16: oligonucleotide attB2 "reverse" primer for amplification of the 200 bp CHS fragment.

SEQ ID No 17: oligonucleotide attB1 "forward" primer used for amplification of 100 bp CHS fragment.

SEQ ID No 18: oligonucleotide attB2 "reverse" primer for amplification of the 100 bp CHS fragment.

SEQ ID No 19: oligonucleotide attB1 "forward" primer used for amplification of 50 bp CHS fragment.

SEQ ID No 20: oligonucleotide attB2 "reverse" primer for amplification of the 50 bp CHS fragment.

SEQ ID No 21: oligonucleotide attB1 "forward" primer for amplification of the 25 bp CHS fragment.

SEQ ID No 22: oligonucleotide attB2 "reverse" primer for the 25 bp fragment.

SEQ ID No 23: nucleotide sequence of the acceptor vector "pHELLSGATE 4"

SEQ ID No 24: nucleotide sequence of the acceptor vector "pHELLSGATE 8"

SEQ ID No 25: nucleotide sequence of the acceptor vector "pHELLSGATE 11"

SEQ ID No 26: nucleotide sequence of the acceptor vector "pHELLSGATE 12"

[097] EXAMPLES

[098] Example 1. Construction of the acceptor vector pHELLSGATE

[099] With the completion of the *Arabidopsis* genome project, the advent of micro-array technology and the ever-increasing investigation into plant metabolic, perception, and response pathways, a rapid targeted way of silencing genes would be of major assistance. The high incidence and degree of silencing in plants transformed with chimeric genes containing simultaneously a sense and antisense nucleotide sequence, as well as a functional intron sequence suggested that such vectors could form the basis of a high-throughput silencing vector. However, one of the major obstacles in using such conventional cloning vectors for a large number of defined genes or a library of undefined genes would be cloning the hairpin arm sequences for each gene in the correct orientations.

[0100] Attempts to clone PCR products of sense and antisense arms together with the appropriately cut vector as a single step four-fragment ligation failed to give efficient or reproducible results. Therefore, a construct (pHELLSGATE) was made to take advantage of Gateway™ (Life Technologies). With this technology, a PCR fragment is generated, bordered with recombination sites (*att*B1 and *att*B2) which is directionally recombined, *in vitro*, into a plasmid containing two sets of suitable recombination sites (*att*P1 and *att*P2 sites) using the commercially available recombination protein preparation.

[0101] The pHELLSGATE vector was designed such that a single PCR product from primers with the appropriate *att*B1 and *att*B2 sites would be recombined into it simultaneously to form the two arms of the hairpin. The *ccd*B gene, which is lethal in standard *E. coli* strains such as DH5α (but not in DB3.1), was placed in the locations to be replaced by the arm sequences,

ensuring that only recombinants containing both arms would be recovered. Placing a chloramphenicol resistance gene within the intron, gives a selection to ensure the retention of the intron in the recombinant plasmid.

[0102] pHELLSGATE comprises the following DNA fragments:

a spectinomycin/streptomycin resistance gene (SEQ ID No 13 from the nucleotide at position 7922 to the nucleotide sequence at 9985);

a right T-DNA border sequence (SEQ ID No 13 from the nucleotide at position 10706 to the nucleotide sequence at 11324);

a CaMV35S promoter (SEQ ID No 13 from the nucleotide at position 11674 to the nucleotide sequence at 13019);

an *att*P1 recombination site (complement of the nucleotide sequence of SEQ ID No 13 from the nucleotide at position 17659 to the nucleotide sequence at 17890);

a *ccd*B selection marker (complement of the nucleotide sequence of SEQ ID No 13 from the nucleotide at position 16855 to the nucleotide at position 17610);

an attP2 recombination site (complement of the nucleotide sequence of SEQ ID No 13 from the nucleotide at position 16319 to the nucleotide at position 16551);

pdk intron2 (SEQ ID No 13 from the nucleotide at position 14660 to the nucleotide at position 16258) flanked by the intron splice site (TACAG*TT (SEQ ID No 13 from the nucleotide at position 16254 to the nucleotide sequence at 16260) and the intron splice site (TG*GTAAG) (SEQ ID No 13 from the nucleotide at position 14660 to the nucleotide sequence at 14667) and comprising a chloramphenicol resistance gene (SEQ ID No 13 from the nucleotide at position 15002 to the nucleotide at position 15661);

an *att*P2 recombination site (SEQ ID No 13 from the nucleotide at position 14387 to the nucleotide at position 14619);

a *ccd*B selection marker (complement of the nucleotide sequence of SEQ ID No 13 from the nucleotide at position 13675 to the nucleotide at position 13980);

an *att*P1 recombination site (SEQ ID No 13 from the nucleotide at position 13048 to the nucleotide at position 13279);

an octopine synthase gene terminator region (SEQ ID No 13 from the nucleotide at position 17922 to the nucleotide sequence at 18687);

a chimeric marker selectable in plants comprising:

a nopaline synthase promoter (SEQ ID No 13 from the nucleotide at position 264 to the nucleotide sequence at 496);

a nptII coding region (SEQ ID No 13 from the nucleotide at position 497 to the nucleotide sequence at 1442); and

a nopaline synthase gene terminator (SEQ ID No 13 from the nucleotide at position 1443 to the nucleotide sequence at 2148);

a left T-DNA border sequence (SEQ ID No 13 from the nucleotide at position 2149 to the nucleotide sequence at 2706);

an origin of replication; and

a kanamycin resistance gene;

[0103] The complete nucleotide sequence of pHELLSGATE is represented in the sequence listing (SEQ ID No 13) and a schematic figure can be found in Fig 3.

[0104] EXAMPLE 2. Use of the pHELLSGATE to convert nucleic acid fragments of interest into dsRNA producing chimeric silencing genes.

[0105] To test the acceptor vector pHELLSGATE, about 400bp, 200bp, 100bp, 50 bp and 25 bp fragments of the *Arabidopsis thaliana* chalcone synthase isomerase coding sequence (SEQ ID No 12) (having respectively the nucleotide sequence of SEQ ID No 12 from the nucleotide at position 83 to the nucleotide at position 482; the nucleotide sequence of SEQ ID No 12 from the nucleotide at position 83 to the nucleotide at position 222; the nucleotide sequence of SEQ ID No 12 from the nucleotide at position 83 to the nucleotide at position 182; the nucleotide

sequence of SEQ ID No 12 from the nucleotide at position 83 to the nucleotide at position 132; and the nucleotide sequence of SEQ ID No 12 from the nucleotide at position 83 to the nucleotide at position 107) were used as nucleic acid fragments of insert for construction of chimeric genes capable of producing dsRNA.

[0106] This gene was chosen because its mutant allele has been reported in *Arabidopsis* to give distinct phenotypes. The CHS tt4(85) EMS mutant (Koornneef, 1990) produces inactive CHS resulting in no anthocyanin pigment in either the stem or seed-coat. Wildtype plants produce the purple-red pigment in both tissues.

[0107] In a first step, the respective fragments were PCR amplified using specific primers further comprising *att*B1 and *att*B2 recombination sites. *Att*B1 and *att*B2 specific primers were purchased from Life Technologies. The 25 and 50 bp fragments flanked by att sites were made by dimerization of the primers.

[0108] The following combinations of primers were used:

For the 400 bp fragment:

Forward primer:

GGGGACAAGTTTGTACAAAAAAGCAGGCTGCACTGCTAACCCTGAGAACCATGTGCTTC

(SEQ ID No 14);

Reverse primer:

GGGGACCACTTTGTACAAGAAAGCTGGGTCGCTTGACGGAAGGACGGAGACCAAGAAGC

(SEQ ID No 15).

For the 200 bp fragment:

Forward primer:

GGGGACAAGTTTGTACAAAAAAGCAGGCTGCACTGCTAACCCTGAGAACCATGTGCTTC

(SEQ ID No 14);

Reverse primer:

GGGGACCACTTTGTACAAGAAAGCTGGGTAGGAGCCATGTAAGCACACATGTGTGGGTT (SEQ ID No 16).

For the 100 bp fragment:

Forward primer:

GGGGACAAGTTTGTACAAAAAAGCAGGCTGCACTGCTAACCCTGAGAACCATGTGCTTCAG GCGGAGTATCCTGACTACTACTTCCGCATCACCAACAGT (SEQ ID No 17);

Reverse primer:

GGGGACCACTTTGTACAAGAAAGCTGGGTAACTTCTCCTTGAGGTCGGTCATGTGTTCACT GTTGGTGATGCGGAAGTAGTAGTCAGGATACTCCGCCTG (SEQ ID No 18).

For the 50 bp fragment:

Forward primer:

GGGGACAAGTTTGTACAAAAAAGCAGGCTGCACTGCTAACCCTGAGAACCATGTGCTTCAG GCGGAGTATCCTGACTAC (SEQ ID No 19);

Reverse primer:

GGGGACCACTTTGTACAAGAAAGCTGGGTGTAGTCAGGATACTCCGCCTGAAGCACATGG TTCTCAGGGTTAGCAGTGC (SEQ ID No 20).

For the 25 bp fragment:

Forward primer:

GGGGACAAGTTTGTACAAAAAAGCAGGCTGCACTGCTAACCCTGAGAACCATGT (SEQ ID No 21);

Reverse primer:

GGGGACCACTTTGTACAAGAAAGCTGGGTACATGGTTCTCAGGGTTAGCAGTGC (SEQ ID No 22).

[0109] PCR amplification and recombination using the GATEWAY™ technology with the commercially available BP Clonase (Life Technologies) were performed according to the

manufacturer's instructions (manual available on http://www.lifetech.com/content.cfm?pageid=2497).

[0110] Bacterial colonies obtained on chloramphenicol-containing plates spread with *E. coli* DH5α bacteria, transformed (by electroporation or by heatshocking RbCl2 treated competent *E. coli* cells) with the *in vitro* recombination reaction were screened. Colonies containing the desired recombinant plasmid were obtained in each case. For the about 400 bp fragment, 24 colonies were screened and 23 contained the desired construct with the 400 bp in inverted repeat, operably linked to the CaMV35S promoter. For the about 200 bp fragment, 36 colonies were screened and 35 contained the desired construct with the 200 bp in inverted repeat, operably linked to the CaMV35S promoter. For the about 50 bp fragment, six colonies were screened and four contained the desired construct with the 50 bp in inverted repeat, operably linked to the CaMV35S promoter. For the 25 bp fragment, six colonies were screened and one contained the desired construct with the 400 bp in inverted repeat, operably linked to the CaMV35S promoter. In a number of cases, the structure was confirmed by sequence analysis.

[0111] These results show that this vector facilitates the rapid, efficient, and simple production of hpRNA (hairpin RNA constructs). pHELLSGATE is a T-DNA vector, with a high-copy-number origin of replication for ease of handling. Recombinant pHELLSGATE constructs can be directly transformed into *Agrobacterium* for transformation of the chimeric construct into plants. This system can be used in high throughput applications.

[0112] EXAMPLE 3. Evaluation of plants comprising the chimeric genes of Example 2.

[0113] The vectors containing the dsRNA producing chimeric constructs with the 400, 200, 100, 50 and 25 nucleotides of chalcone synthase in inverted repeat (Example 2) were introduced into *Agrobacterium tumefaciens* strain AGL1, GV3101 or LBA4404 either by electroporation or tri-parental mating.

[0114] Transgenic *Arabidopsis* lines are obtained by transformation with these *Agrobacteria* using the dipping method of Clough and Bent (1998).

[0115] Chalcone synthase activity is monitored by visual observation of stem and leaf color (normally in plants grown under high light, and by unaided or microscope assisted visual observation of seed-coat color.

[0116] Most of the transgenic lines transformed with the above-mentioned CHS silencing constructs show pronounced silencing. The seed color of most of these lines is virtually indistinguishable from seed of the tt4(85) mutant to the naked eye. Examination of the seed under a light microscope reveals that the degree of pigmentation is generally uniform in the cells of the coat of an individual seed, and among seeds of the same line.

[0117] EXAMPLE 4. Construction of the acceptor vectors pHELLSGATE 4, pHELLSGATE 8, pHELLSGATE 11 and pHELLSGATE 12.

[0118] pHELLSGATE 4 was made by excising the DNA fragment comprising the *pdk* intron and chloramphenicol resistance gene from pHellsgate (Example 1) with *Hin*dIII and *Eco*RI and replacing it with a HindIII/EcorRI DNA fragment containing only the *pdk* intron. The complete nucleotide sequence of pHELLSGATE 4 is represented in the sequence listing (SEQ ID No 23).

[0119] pHellsgate 8 was made by PCR amplification using pHellsgate DNA as a template and oligonucleotides with the sequence 5'GGGCTCGAGACAAGTTTGTACAAAAAAGCTG 3' and 5'GGCTCGAGACCACTTTGTACAAGAAAGC 3' as primers. These primers modify the attP sites within pHellsgate to attR sites. The resulting fragment was sequenced and inserted into the *Xho*I site of a vector upstream of a DNA fragment containing the *pdk* intron fragment. Similarly an *Xba*I/*Xba*I fragment amplified with the oligonucleotides 5'GGGTCTAGACAAGTTTGTACAAAAAAGCTG 3' and 5'GGGTCTAGACCACTTTGTACAAGAAAGC 3' as primers and pHEllSGATE as template DNA to modify the attP sites of this cassette to attR sites. This fragment was sequenced and inserted

into the *Xba*I site of the intermediate described above downstream of the *pdk* intron. The complete nucleotide sequence of pHELLSGATE 8 is represented in the sequence listing (SEQ ID No 24) and a schematic figure can be found in Fig 4.

[0120] pHELLSGATE 11 is similar to pHELLSGATE 8 except that the pdk intron has been engineered to contain a branching point in the complementary strand such that splicing of the intron is independent of its orientation (a so-called "two-way intron"). The complete nucleotide sequence of pHELLSGATE 11 is represented in the sequence listing (SEQ ID No 25) and a schematic representation thereof can be found in Fig 4.

[0121] pHELLSGATE 12 is also similar to pHELLSGATE 8 except that the pdk intron has been duplicated as an inverted repeat. The complete nucleotide sequence of pHELLSGATE 12 is represented in the sequence listing (SEQ ID No 26) and a schematic representation thereof can be found in Fig 4.

[0122] Example 5. Use of the different pHELLSGATE vectors to generate dsRNA chimeric silencing genes targeted towards three different model target genes.

[0123] The efficiency in gene silencing of the different pHELLSGATE vectors was tested by inserting fragments of three target genes: Flowering locus C (FLC); Ethylene insensitive 2 (EIN2); and Phytoene desaturase (PDC). For FLC a 390 bp fragment was used (from the nucleotide at position 303 to the nucleotide at position 692 of the nucleotide sequence available as Genbank Accession Nr AF116527). For EIN2 a 580 bp fragment was used (from the nucleotide at position 541 to the nucleotide at position 1120 of the nucleotide sequence available as Genbank Accession Nr AF141203). For PDS a 432 bp fragment was used (from the nucleotide at position 1027 to the nucleotide at position 1458 of the nucleotide sequence available as Genbank Accession Nr L16237). Genes of interest were amplified using gene specific primers with either a 5' attB1 extension (GGGGACAAGTTTGTACAAAAAAGCAGGCT) or an attB2 extension (GGGACCACTTTGTACAAGAAAGCTGGGT) using F1 *Taq* DNA

polymerase (Fisher Biotec, Subiaco, WA, Australia) according to the manufacturer's protocol. PCR products were precipitated by adding 3 volumes TE and two volumes 30% (w/v) PEG 3000, 30mM $MgCl_2$ and centrifuging at 13000 g for 15 minutes.

[0124] Recombination reaction of PCR products with either pDONR201 (Invitrogen, Groningen, The Netherlands) or pHellsgate 4 were carried out in a total volume of 10 μL with 2 μL BP clonase buffer (Invitrogen), 1-2 μL PCR product 150 ng plasmid vector and 2 μL BP clonase (Invitrogen). The reaction was incubated at room temperature (25°C) for 1 h to overnight. After the incubation, 1 μL proteinase K (2 μg/μL; Invitrogen) was added and incubated for 10 min at 37°C. 1-2 μL of the mix was used to transform DH5α; colonies were selected on the appropriate antibiotics. Clones were checked either by digestion of DNA minipreps or PCR. Recombination reactions from pDONR201 clones to pHellsgate 8, 11 or 12 were carried out in 10 μL total volume with 2 μL LR clonase buffer (Invitrogen), 2 μL pDONR201 clone (approximately 150 ng), 300 ng pHellsgate 8, 11 or 12 and 2 μL LR clonase (Invitrogen). The reaction was incubated overnight at room temperature, proteinase-treated and used to transform E. coli DH5α as for the BP clonase reaction.

[0125] Transformation of *Arabidopsis* was performed according to via the floral dip method (Clough and Bent, 1998). Plants were selected on agar solidified MS media supplemented with 100 mg/l timentin and 50 mg/l kanamycin. For *FLC* and *PDS* constructs, the C24 ecotype was used; for *EIN2* constructs, Landsberg *erecta* was used. For scoring of *EIN2* phenotypes, transformed T1 plants were transferred to MS media containing 50 μM 1-aminocyclopropane-1-carboxylic acid (ACC) together with homozygous *EIN2*-silenced lines and wild type Landberg erecta plants. T1 *FLC* hpRNA plants were scored by transferring to MS plates and scoring days to flower or rosette leaves at flowering compared to C24 wild type plants and *flc* mutant lines. T1 PDS hpRNA plants were scored by looking at bleaching of the leaves. The results of the analysis of plants transformed with the different pHELLSGATE vectors are shown in Table 1.

[0126] All plants transformed with pHellsgate 4-FLC and pHellsgate 8-FLC flowered significantly earlier than wildtype C24 and in both cases plants flowering with the same number of rosette leaves as the *flc-20* line (carrying a stable Ds insertion in the first intron of the *FLC* gene) were observed. There was no clear difference in rosette leaves at flowering between the sets of plants transformed with the pHellsgate 4-FLC and pHellsgate 8-FLC constructs.

[0127] A difference in the effectiveness of the pHellsgate 4-EIN2 and pHellsgate 8-EIN2 plants was observed. Of 36 transformants for pHG4-EIN2, there were no plants with an observable ACC-resistant phenotype under the conditions used for this experiment, whereas eight of the 11 plants carrying the pHG8-EIN2 transgene showed some degree of ACC-resistance. The extent to which the pHG8-EIN2 plants were resistant to ACC was variable indicating that the severity of silencing varies between transformants.

[0128] The great majority of plants carrying pHG4-PDS and pHG8-PDS showed a phenotype consistent with the loss of photoprotection due to the absence of carotenoids. The weakest phenotype was a bleaching of the cotyledons, with the true leaves not bleaching at any stage in the life cycle. The bleached cotyledon phenotype was only seen in plants transformed with *PDS* hpRNA constructs; we confirmed that the plants with this phenotype also contained the *PDS* hpRNA construct (data not shown) strongly suggesting that this phenotype is due to *PDS* silencing and not bleaching from the kanamycin selection. Plants transformed with the pHellsgate 4-PDS construct gave only this weak bleached cotyledon phenotype. In contrast, the five of the pHellsgate 8-PDS plants had the weak phenotype and three showed a stronger phenotype with extensive or complete bleaching of the true leaves.

[0129] Table 1

Construct	Test genes	T1 plants	Rate of silencing
HELLSGATE 4	FLC EIN2 PDS	13 36 12	12 0 11
HELLSGATE 8	FLC EIN2 PDS	6 11 9	6 8 8
HELLSGATE 11	FLC EIN2 PDS	2 30 11	2 11 11
HELLSGATE 11 (intervening region in inverse orientation)	FLC EIN2 PDS	8	6
HELLSGATE 12	FLC EIN2 PDS	13 26	11 12
HELLSGATE 12 (intervening region in inverse orientation)	FLC EIN2 PDS CHS	9 5	8 2

[0130] References

[0131] An *et al.* (1996) *The Plant Cell* 8, 15-30

[0132] AzpiroLeehan and Feldmann (1997) *Trends Genet.* 13: 152-156

[0133] Clough and Bent (1998) *Plant J.* 16: 735-743

[0134] Fire *et al.* (1998) *Nature* 391: 806-811

[0135] Fromm *et al.* (1990) *Bio/Technology* 8: 833

[0136] Gordon-Kamm *et al.* (1990) *The Plant Cell* 2: 603

[0137] Hamilton *et al.* (1998) *Plant J.* 15: 737-746

[0138] Hoess *et al.* (1986) *Nucl. Acids Res.* 14: 2287

[0139] Hudspeth *et al.* (1989) *Plant Mol. Biol.* 12: 579-589

[0140] Keil *et al.* (1989) *EMBO J.* 8: 1323-1330

[0141] Keller *et al.* (1988) *EMBO J.* 7: 3625-3633

[0142] Keller *et al.* (1989) *Genes & Devel.* 3: 1639-1646

[0143] Koornneef (1990) *Theor. Appl. Gen.* 80: 852-857

[0144] Landy (1993) Current Opinions in Genetics and Development 3: 699-707

[0145] Landy (1989) *Ann. Rev. Biochem.* 58: 913

[0146] Martienssen (1998) *Proc. Natl. Acad. Sci. USA* 95: 2021-2026

[0147] Meyer *et al.* (1987) *Nature* 330: 677

[0148] Needleman and Wunsch (1970) *J. Mol. Biol.* 48: 443-453

[0149] Peleman *et al.* (1989) *Gene* 85: 359-369

[0150] Ross-MacDonald *et al.* (1999) *Nature* 402: 413-418

[0151] Smith *et al.* (2000) *Nature* 407: 319-320

[0152] Wagner and Sun (1998) *Nature* 391: 744-745

[0153] Waterhouse *et al* (1998) *Proc. Natl. Acad. Sci. USA* 95: 13959-13964